

!!NA MULTIPLE ALIGNMENT 1.0  
Pileup of: us\*

Symbol comparison table: GenRunData:pileupdna.cmp CompCheck: 6876

GapWeight: 5  
GapLengthWeight: 0

pileup.msf MSF: 2307 Type: N October 28, 2003 09:08 Check: 4174 ..

Name: us-09-872-523-6 Len: 2307 Check: 7087 Weight: 1.00  
Name: us-09-872-523-73 Len: 2307 Check: 7818 Weight: 1.00  
Name: us-09-872-523-74 Len: 2307 Check: 6841 Weight: 1.00  
Name: us-09-872-523-75 Len: 2307 Check: 7223 Weight: 1.00  
Name: us-09-872-523-78 Len: 2307 Check: 5205 Weight: 1.00

//

```
us-09-872-523-6 1 ~~~~~ 50
us-09-872-523-73 ~~~~~
us-09-872-523-74 ~~~~~
us-09-872-523-75 ~~~~~
us-09-872-523-78 atgtcaaat tagtatttt gtgtcttcgt ttgtctaca atacagtcag
51 ~~~~~ 100
us-09-872-523-6 ~~~~~
us-09-872-523-73 ~~~~~
us-09-872-523-74 ~~~~~
us-09-872-523-75 ~~~~~
us-09-872-523-78 ttctgcacaga ttctgtcttg ggtctgcagat taagtgcgat ttgtatgaag
101 ~~~~~ 150
us-09-872-523-6 ~~~~~
us-09-872-523-73 ~~~~~
us-09-872-523-74 ~~~~~
us-09-872-523-75 ~~~~~
us-09-872-523-78 tgttccaatt aacagtgctg catttggaag acgttgacaa tacttttgg
151 ~~~~~ 200
us-09-872-523-6 ~~~~~
us-09-872-523-73 ~~~~~
us-09-872-523-74 ~~~~~
us-09-872-523-75 ~~~~~
us-09-872-523-78 gatcgcgatg aagacatcac tggacgtatg actatgttgg ctgcagaagaa
201 ~~~~~ 250
us-09-872-523-6 ~~~~~
us-09-872-523-73 ~~~~~
us-09-872-523-74 ~~~~~
us-09-872-523-75 ~~~~~
us-09-872-523-78 aatatcttcc tatcagagcg gccatcatgg atttgaatt ggaagctcg
251 ~~~~~ 300
us-09-872-523-6 ~~~~~
us-09-872-523-73 ~~~~~
us-09-872-523-74 ~~~~~
us-09-872-523-75 ~~~~~
us-09-872-523-78 agccttatgg gtgttcttg cacaattgca cgaagaatgg aaatttctgc
301 ~~~~~ 350
us-09-872-523-6 ~~~~~
us-09-872-523-73 ~~~~~
us-09-872-523-74 ~~~~~
us-09-872-523-75 ~~~~~
us-09-872-523-78 gagtataggc acggtttgag tagcacagct ggaatccaatg ggttggagta
351 ~~~~~ 400
us-09-872-523-6 ~~~~~
us-09-872-523-73 ~~~~~
us-09-872-523-74 ~~~~~
us-09-872-523-75 ~~~~~
us-09-872-523-78 tatggagttac acgttgaatt tgacgaagcg ctgcagaata tcaaatcaaa
```

```
us-09-872-523-6 401 ~~~~~ 450
us-09-872-523-73 ~~~~~
us-09-872-523-74 ~~~~~
us-09-872-523-75 ~~~~~
us-09-872-523-78 tctgaactaac ttccaattcc aataacat ctctccaatt acgtttttaa
451 ~~~~~ 500
us-09-872-523-6 ~~~~~
us-09-872-523-73 ~~~~~
us-09-872-523-74 ~~~~~
us-09-872-523-75 ~~~~~
us-09-872-523-78 ccagttttaa ttccagatgt ctgatattct gaaattgtgc agagctcaaca
501 ~~~~~ 550
us-09-872-523-6 ~~~~~
us-09-872-523-73 ~~~~~
us-09-872-523-74 ~~~~~
us-09-872-523-75 ~~~~~
us-09-872-523-78 aaaaatcggc cagaaatcgc gataagacct acttgcggga atccatttta
551 ~~~~~ 600
us-09-872-523-6 ~~~~~
us-09-872-523-73 ~~~~~
us-09-872-523-74 ~~~~~
us-09-872-523-75 ~~~~~
us-09-872-523-78 catcagttcg agaaaggaaa aactctcttc attccagttg aagcatctcaa
601 ~~~~~ 650
us-09-872-523-6 ~~~~~
us-09-872-523-73 ~~~~~
us-09-872-523-74 ~~~~~
us-09-872-523-75 ~~~~~
us-09-872-523-78 tctgaacct acagtttaatt ttaacgaatg cgtggaaggaa ggaattatc.
651 ~~~~~ 700
us-09-872-523-6 ~~~~~
us-09-872-523-73 ~~~~~
us-09-872-523-74 ~~~~~
us-09-872-523-75 ~~~~~
us-09-872-523-78 tctgaacct acagtttaatt ttaacgaatg cgtggaaggaa ggaattatc.
651 ~~~~~ 700
us-09-872-523-6 ~~~~~
us-09-872-523-73 ~~~~~
us-09-872-523-74 ~~~~~
us-09-872-523-75 ~~~~~
us-09-872-523-78 tgaattcata ttgtctgtaa atcgttttta aaatacaatt ttgttgatc
701 ~~~~~ 750
us-09-872-523-6 ~~~~~
us-09-872-523-73 ~~~~~
us-09-872-523-74 ~~~~~
us-09-872-523-75 ~~~~~
us-09-872-523-78 gaaacagtcg tccatgatata tgacaagaac tgcgattcga ttcaagtcag
751 ~~~~~ 800
us-09-872-523-6 ~~~~~
us-09-872-523-73 ~~~~~
us-09-872-523-74 ~~~~~
us-09-872-523-75 ~~~~~
us-09-872-523-78 atgtgtttgca cgaattgaaa aagtttgcgg atacagaagt cctggctcaat
801 ~~~~~ 850
us-09-872-523-6 ~~~~~
us-09-872-523-73 ~~~~~
us-09-872-523-74 ~~~~~
us-09-872-523-75 ~~~~~
us-09-872-523-78 ttaacggagc tgacacgaaa tttrgtctca atattttac ggaagatag
851 ~~~~~ 900
us-09-872-523-6 ~~~~~
us-09-872-523-73 ~~~~~
us-09-872-523-74 ~~~~~
us-09-872-523-75 ~~~~~
us-09-872-523-78 ttgtgttttg caaa..... ttgtgttttg caaa..... ttgtgttttg caaa.....
901 ~~~~~ 950
us-09-872-523-6 ~~~~~
us-09-872-523-73 ~~~~~
us-09-872-523-74 ~~~~~
us-09-872-523-75 ~~~~~
us-09-872-523-78 ttgtgttttg caaa..... ttgtgttttg caaa..... ttgtgttttg caaa.....
901 ~~~~~ 950
```

us-09-872-523-73	.....	.....	.....	.....	us-09-872-523-78	ttgtagatgac	gagggagaagct	tcttcatact	tcctgtttgga	tttgccgcgcg
us-09-872-523-74	.....	.....	.....	.....	1451	lcaatgagata	lcaactaat	gcgaaaagag	aatatatitga	gcacacaaat
us-09-872-523-75	.....	.....	.....	.....	us-09-872-523-76	lcaatgagata	lcaactaat	gcgaaaagag	aatatatitga	gcacacaaat
us-09-872-523-78	taataatagc	ttctgttaca	taaatitcta	gagagaacatc	gratlaanaac	us-09-872-523-77	lcaatgagata	lcaactaat	gcgaaaagag	aatatatitga
us-09-872-523-76	951	.....	.....	.....	1000	gcgcgcgaag	.....	.....	.....	.....
us-09-872-523-73	.....	.....	.....	.....	gcgcgcgaag	.....	.....	.....	.....	.....
us-09-872-523-74	.....	.....	.....	.....	gcgcgcgaag	.....	.....	.....	.....	.....
us-09-872-523-75	.....	.....	.....	.....	gcgcgcgaag	.....	.....	.....	.....	.....
us-09-872-523-78	ttcgaanaact	ttgtataraa	gtaanaattg	aaactttcag	cgccgcgaatg	us-09-872-523-6	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-73	1001	agagatccca	atatgagata	aatvtatat	gctccgcgcg	ttgcaatcaa	1050	1501	aaatttgc	aaagcataaa
us-09-872-523-74	agagatccca	atatgagata	aatvtatat	gctccgcgcg	ttgcaatcaa	us-09-872-523-73	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-75	agagatccca	atatgagata	aatvtatat	gctccgcgcg	ttgcaatcaa	us-09-872-523-74	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-78	agagatccca	atatgagata	aatvtatat	gctccgcgcg	ttgcaatcaa	us-09-872-523-75	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-6	1051	cgaagaatac	caaaatgata	tgtranaat	.....	1100	1511	1601	aaatttgc	aaagcataaa
us-09-872-523-73	cgaagaatac	caaaatgata	tgtranaat	.....	.....	us-09-872-523-73	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-74	cgaagaatac	caaaatgata	tgtranaat	.....	.....	us-09-872-523-74	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-75	cgaagaatac	caaaatgata	tgtranaat	.....	.....	us-09-872-523-75	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-78	cgaagaatac	caaaatgata	tgtranaat	.....	.....	us-09-872-523-78	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-6	1101	.....	.....	.....	1150	1611	1650	1700	aaatttgc	aaagcataaa
us-09-872-523-73	.....	.....	.....	.....	.....	us-09-872-523-73	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-74	.....	.....	.....	.....	.....	us-09-872-523-74	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-75	.....	.....	.....	.....	.....	us-09-872-523-75	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-78	tttcggaatt	latgttaata	lcatctcaca	acttcaaat	ttgatitgag	us-09-872-523-6	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-6	1151	gcggaatacgt	cgccgcaaat	tcgcgtgttc	caaaatlcga	tgaaagggaag	1200	1701	aaatttgc	aaagcataaa
us-09-872-523-73	gcggaatacgt	cgccgcaaat	tcgcgtgttc	caaaatlcga	tgaaagggaag	us-09-872-523-73	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-74	gcggaatacgt	cgccgcaaat	tcgcgtgttc	caaaatlcga	tgaaagggaag	us-09-872-523-74	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-75	gcggaatacgt	cgccgcaaat	tcgcgtgttc	caaaatlcga	tgaaagggaag	us-09-872-523-75	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-78	gcggaatacgt	cgccgcaaat	tcgcgtgttc	caaaatlcga	tgaaagggaag	us-09-872-523-78	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-6	1201	gcttcctcta	gcaagatcgc	tttcaaaat	ggacaacgic	ttgaacttat	1250	1751	aaatttgc	aaagcataaa
us-09-872-523-73	gcttcctcta	gcaagatcgc	tttcaaaat	ggacaacgic	ttgaacttat	us-09-872-523-73	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-74	gcttcctcta	gcaagatcgc	tttcaaaat	ggacaacgic	ttgaacttat	us-09-872-523-74	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-75	gcttcctcta	gcaagatcgc	tttcaaaat	ggacaacgic	ttgaacttat	us-09-872-523-75	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-78	gcttcctcta	gcaagatcgc	tttcaaaat	ggacaacgic	ttgaacttat	us-09-872-523-78	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-6	1251	aaatattc	aatctcagc	aaatccgcgt	agccgcgaat	caagaaatct	1300	1801	aaatttgc	aaagcataaa
us-09-872-523-73	aaatattc	aatctcagc	aaatccgcgt	agccgcgaat	caagaaatct	us-09-872-523-73	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-74	aaatattc	aatctcagc	aaatccgcgt	agccgcgaat	caagaaatct	us-09-872-523-74	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-75	aaatattc	aatctcagc	aaatccgcgt	agccgcgaat	caagaaatct	us-09-872-523-75	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-78	aaatattc	aatctcagc	aaatccgcgt	agccgcgaat	caagaaatct	us-09-872-523-78	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-6	1301	gtggagacgc	aatgaatgta	ttctacacaa	agaaaagact	ttccggaatcg	1350	1850	aaatttgc	aaagcataaa
us-09-872-523-73	gtggagacgc	aatgaatgta	ttctacacaa	agaaaagact	ttccggaatcg	us-09-872-523-73	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-74	gtggagacgc	aatgaatgta	ttctacacaa	agaaaagact	ttccggaatcg	us-09-872-523-74	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-75	gtggagacgc	aatgaatgta	ttctacacaa	agaaaagact	ttccggaatcg	us-09-872-523-75	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-78	gtggagacgc	aatgaatgta	ttctacacaa	agaaaagact	ttccggaatcg	us-09-872-523-78	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-6	1351	cttcacagat	cagatgacga	cagacaagtc	tttagctctg	gattcccaata	1400	1900	aaatttgc	aaagcataaa
us-09-872-523-73	cttcacagat	cagatgacga	cagacaagtc	tttagctctg	gattcccaata	us-09-872-523-73	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-74	cttcacagat	cagatgacga	cagacaagtc	tttagctctg	gattcccaata	us-09-872-523-74	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-75	cttcacagat	cagatgacga	cagacaagtc	tttagctctg	gattcccaata	us-09-872-523-75	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-78	cttcacagat	cagatgacga	cagacaagtc	tttagctctg	gattcccaata	us-09-872-523-78	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-6	1401	ttgtagatgac	gagggagaagct	tcttcatact	tcctgtttgga	tttgccgcgcg	1450	1950	aaatttgc	aaagcataaa
us-09-872-523-73	ttgtagatgac	gagggagaagct	tcttcatact	tcctgtttgga	tttgccgcgcg	us-09-872-523-73	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-74	ttgtagatgac	gagggagaagct	tcttcatact	tcctgtttgga	tttgccgcgcg	us-09-872-523-74	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-75	ttgtagatgac	gagggagaagct	tcttcatact	tcctgtttgga	tttgccgcgcg	us-09-872-523-75	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-78	ttgtagatgac	gagggagaagct	tcttcatact	tcctgtttgga	tttgccgcgcg	us-09-872-523-78	aaatttgc	aaagcataaa	aaatggagaa	aatccaaagat
us-09-872-523-6	1450	.....	.....	.....	.....	.....	2000	.....	.....	.....

```
us-09-872-523-6 attctgattt ccaacg... ..
us-09-872-523-73 attctgattt ccaacg... ..
us-09-872-523-74 attctgattt ccaacg... ..
us-09-872-523-75 attctgattt ccaacg... ..
us-09-872-523-78 attctgattt ccaacgggt tgtttatat cgtttagat tgtttcata
2001
us-09-872-523-6 ..gtagga ctacgcttg 2050
us-09-872-523-73 ..gtagga ctacgcttg
us-09-872-523-74 ..gtagga ctacgcttg
us-09-872-523-75 ..gtagga ctacgcttg
us-09-872-523-78 taagttagga ctacgcttg
2051
us-09-872-523-6 aagctgctga catgttgaa aatcagttta ttgtccagc tacagtgaaa 2100
us-09-872-523-73 aagctgctga catgttgaa aatcagttta ttgtccagc tacagtgaaa
us-09-872-523-74 aagctgctga catgttgaa aatcagttta ttgtccagc tacagtgaaa
us-09-872-523-75 aagctgctga catgttgaa aatcagttta ttgtccagc tacagtgaaa
us-09-872-523-78 aagctgctga catgttgaa aatcagttta ttgtccagc tacagtgaaa
2101
us-09-872-523-6 tcaattcatg gaagactgat aaatgtcaat ttcgacggct gggatgaaga 2150
us-09-872-523-73 tcaattcatg gaagactgat aaatgtcaat ttcgacggct gggatgaaga
us-09-872-523-74 tcaattcatg gaagactgat aaatgtcaat ttcgacggct gggatgaaga
us-09-872-523-75 tcaattcatg gaagactgat aaatgtcaat ttcgacggct gggatgaaga
us-09-872-523-78 tcaattcatg gaagactgat aaatgtcaat ttcgacggct gggatgaaga
2151
us-09-872-523-6 atttgatgaa ctgtatgatg tggg... .. 2200
us-09-872-523-73 atttgatgaa ctgtatgatg tggg... ..
us-09-872-523-74 atttgatgaa ctgtatgatg tggg... ..
us-09-872-523-75 atttgatgaa ctgtatgatg tggg... ..
us-09-872-523-78 atttgatgaa ctgtatgatg tgggagtagt ttatcagac cgaagacat
2201
us-09-872-523-6 ..ccatgatat tctaccgata 2250
us-09-872-523-73 ..ccatgatat tctaccgata
us-09-872-523-74 ..ccatgatat tctaccgata
us-09-872-523-75 ..ccatgatat tctaccgata
us-09-872-523-78 ttcttcaatg aaatcttat cattcagct cccatgatat tctaccgata
2251
us-09-872-523-6 ggaatggttg aagcgacag ttatgttcta caacctcoga aaaagtaaa 2300
us-09-872-523-73 ggaatggttg aagcgacag ttatgttcta caacctcoga aaaagtaaa
us-09-872-523-74 ggaatggttg aagcgacag ttatgttcta caacctcoga aaaagtaaa
us-09-872-523-75 ggaatggttg aagcgacag ttatgttcta caacctcoga aaaagtaaa
us-09-872-523-78 ggaatggttg aagcgacag ttatgttcta caacctcoga aaaagtaaa
2301
us-09-872-523-6 ctattga
us-09-872-523-73 ctattga
us-09-872-523-74 ctattga
us-09-872-523-75 ctattga
us-09-872-523-78 ctattga
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 28, 2003, 09:42:00 : Search time 13 seconds

(without alignments)  
2.368 Million cell updates/sec

Title: us-09-872-523-5  
Perfect score: 2656  
Sequence: 1 MSELKIVRANKSDRKLK.....PIGCEAHSYVLQPPKXNY 498

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1 segs, 30911 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1 summaries

Command line parameters:  
-MODL=frame+ p2n model -DEV=soft -Q=us-09-872-523-5 -DB=19-apr-1996-z71266  
-SUFFIX=ptco -OUT=align5\_1996 -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0 -UNITS=bits  
-START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=1 -DOCALLIGN=200  
-THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=1 -MODE=LOCAL -OUTFMT=ptco  
-NORM=ext -HEAPSIZ=500 -MINLEN=6 -MAXLEN=2000000000 -NCPU=6 -NO\_XLPHY  
-NEG SCORES=0 -LONGLOG -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: 19-apr-1996-z71266.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	2520.5	94.9	30911	1 CER06C7	ALIGNMENTS ACCESSION:Z71266

#### ALIGNMENTS

RESULT 1  
CER06C7/c CER06C7 30911 bp DNA linear INV 19-APR-1996  
DEFINITION Caenorhabditis elegans cosmid R06C7.  
ACCESSION Z71266  
VERSION Z71266  
KEYWORDS adenylosuccinate lyase; myosin heavy chain; Rat TOAD-64 protein  
like; serine/threonine-protein kinase; zinc finger protein.  
SOURCE Caenorhabditis elegans  
ORGANISM Caenorhabditis elegans

REFERENCE  
1 (bases 1 to 30911)  
Rhabditidae; Caenorhabditis.  
AUTHORS Gardner, A.  
TITLE Direct Submission

#### JOURNAL

#### REFERENCE AUTHORS

Submitted (19-APR-1996) Nematode Sequencing Project, Sanger Centre,  
Hinxton, Cambridge CB10 1HQ, England and Department of Genetics,  
Washington University, St. Louis, MO 63110, USA. E-mail:  
jes@sanger.ac.uk or twenematode.wustl.edu  
2 (bases 1 to 30911)  
Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Betks, M.,  
Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J.,  
Coulson, A., Cratton, M., Dear, S., Du, Z., Durbin, R., Favello, A.,  
Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M.,  
Johnston, L., Jones, M., Kerhaw, J., Kirsten, J., Laister, N.,  
Latreille, P., Lightning, J., Lloyd, C., McMurray, A., Mortimore, B.,  
O'Callaghan, M., Parsons, J., Percy, C., Rifken, L., Roopra, A.,  
Saunders, D., Showkeen, R., Smaldon, N., Smith, A., Sonhammer, E.,  
Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M.,  
Vaughan, K., Waterston, R., Watson, A., Weinstock, L.,  
Wilkinson-Sproat, J., and Wohldman, P.  
2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
elegans

#### TITLE JOURNAL MEDLINE COMMENT

Current sequence finishing criteria for the C. elegans genome  
sequencing consortium are that all bases are either sequenced  
unambiguously on both strands, or on a single strand with both a  
dye primer and dye terminator reaction, from distinct subclones.  
Exceptions are indicated by an explicit note.  
IMPORTANT: This sequence is NOT necessarily the entire insert of  
clone R06C7. It may be shorter because we only sequence  
overlapping sections once, or longer because we arrange for a small  
overlap between neighbouring subclones.  
The true left end of clone R06C7 is at 1 in this sequence. The true  
left end of clone F21c3 is at 30808 in this sequence. Coding the  
program genefinder (P. Green, ms in preparation), and other  
available information.  
The end of this sequence (30808..30911) overlaps with the start of  
sequence CEP21C3.

#### FEATURES SOURCE

#### CDS

1..30911  
/organism="Caenorhabditis elegans"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:6239"  
/chromosome="1"  
/clone="R06C7"  
complement(join(1315..1480,1528..1715,1747..1783,  
1834..2156))  
/product="1"  
/codon\_start=1  
/product="R06C7.2"  
/protein\_id="1279328"  
/db\_xref="GI:1279328"

#### CDS

join(3063..3227,3274..3854,3899..4631,4680..4855,  
4899..5074,5117..5463,5510..5749,5790..6097,6143..6254)  
/note="protein predicted using GeneFinder; CDS EST  
yk25b2.5 comes from this gene; cDNA EST yk3446.5 comes  
from this gene; cDNA EST yk31a12.5 comes from this gene;  
cDNA EST yk31694.5 comes from this gene; cDNA EST yk2191.5  
comes from this gene; cDNA EST yk31a11.5 comes from this  
gene; cDNA EST yk31a4.5 comes from this gene; cDNA EST  
yk31a4.3 comes from this gene; cDNA EST yk2191.3 comes  
from this gene; cDNA EST yk25b2.3 comes from this gene;  
cDNA EST yk31a12.3 comes from this gene; cDNA EST yk31694.3  
comes from this gene; cDNA EST yk3444.3 comes from this  
gene"  
/codon\_start=1  
/product="R06C7.1"  
/protein\_id="1279326"  
/db\_xref="GI:1279326"  
/translation="MSPPPPQPPHPPMPMPVTPAPPGAMTPMPVPADAKLHSTONK  
DACIKRLQQLNVEDGAKMYKKPTEPGMGKRPVDIGTNVGEIETVKETTVRRFVWHAKA

**THIS PAGE BLANK (USPTO)**

```

Db      18171 GAAAAAGTTGCGGATACAGAGCTTCGCTCAGTTTATCGAGCTGACAGAAATTTTGG 18112
QY      105 LeuAsn11eLeuSerAspSpmecPheGlyLeuAla----- 116
Db      18111 CTCAAATATTTTTCGAGCAGATATGTTGGTTGGCAAGTAAGTTGACGCTCAGCTCTT 18052
QY      116 ----- 116
Db      18051 TCTACTATTCTAAATAAATATAGTTCCTCTACATAAAATTCAGAGAACATTCGATATTA 17992
QY      117 -----AsnAla1aMetSerAsp 122
Db      17991 AAACCTCGAAACATTGTATTAATAGTAATTAAGTAACATTTCAGCGCCCAATGAGTGAT 17932
QY      123 ProAsnMetAspLyse11eValTyrAlaProProluAla11eAsnGluGluTyrGlnAsn 142
Db      17931 CCCAAATATGATTAATAATTTATATATGCTCCGCCGCTTGCAATCAACGAAATATCCAAAT 17872
QY      143 AspMetValAsnTyrValAsn----- 149
Db      17871 GATATGTTAAATTTATGTAATA--TGTAAGTTGTTTTTTCGCAATTTATGTAATATCATC 17813
QY      150 -----AsnGly11eAspGlyGlu11eValGlyGlnThrSerLeuSerProLys 165
Db      17812 TCACAACTTCAGAAATTGCATTTATGCGCAAAATCGTCGSCCAAACTTCGCTGCTCCAAA 17753
QY      166 PheAspGluGlyLysAlaLeuLeuSerLysHisArgPheLysValGlyGlnArgLeuGlu 185
Db      17752 TTCGATGAAGGAGAGGCTCTCTCTAAGCAACGATCGTTTCAAGTTGGAACAGCTCTTGA 17693
QY      186 LeuLeuAsnTyrSerAsnSerThrglu11eArgValAlaArg11eGlnGlu11eCysGly 205
Db      17692 CTATTAAATTTATTCATCTCTAAGTAAATCCGCTAGCCGCAATTCAGAAATATGTCGA 17633
QY      206 ArgArgMetAsnValSer11eThrLysLysAspPheProGluSerLeuProAspAlaAsp 225
Db      17632 CGACCAATGAATGTATCTATCTCAAAAGAAAGACTTCCGCAATCGCTTCAGATGCAGAT 17573
QY      226 AspAspArgGlnValPheSerSerGlySerGlnTyrTrp11eAspGluGlySerPhePhe 245
Db      17572 GACGACAGCAAGCTCTTAGCTCTGATCTCAATATGATAGACGAGGAAAGCTTCTTC 17513
QY      246 11ePheProValGlyPheAlaAlaValAsnGlyTyrGlnLeuAsnAlaLysLysGluTyr 265
Db      17512 ATATTTCCTGCTGATTTGCACAGTCATTCATCAACTAAATCCGAAAAAGAAATTT 17453
QY      266 11eGluHisThrAsnLys11eAlaGlnAla11eLysAsnGlyGluAsnProArgTyrAsp 285
Db      17452 ATTGAGCACACAAATTAATTTGCTCAAGCAATTAATAATGAGAAATCCAGATATGAC 17393
QY      286 SerAspAspValThrPheAspGlnLeuAlaLysAspPro11eAspProMet11eTrpArg 305
Db      17392 TCAGACGACGTCACATTTGATCAATTAAGCAAAAGATCCCAATTGATCCCATGATTTGAGA 17333
QY      306 LysValLysValGlyGlnLysPheGluLeu11eAspProLeuAlaGlnGlnPheAsnAsn 325
Db      17332 AAAGTTAAGGTTGCAAAAGTTTGAAGTCATCGAACCCCTTGCTCAGCAATTCATAATAC 17273
QY      326 LeuHisValAlaSer11eLeuLysPheCysLysThrGluGlyTyrLeu11eValGlyMet 345
Db      17272 CTCACAGTCGCTTCGATTTCTCAAAATTTTGCAAAACGAAAGATATCTTATTTGGGAATG 17213
QY      346 AspGlyProAspAlaLeuGluLysSerPhePro11eHis11eAsnAsnThrPheMetPhe 365
Db      17212 GATGTCAGAGTCACTTGAAACAGTTTCTATTCATATCAATTAATACATTTATGTTTC 17153
QY      366 ProValGlyTyrAlaGluLysTyrAsnLeuGluLeuVal11eProProAspGluPheLysGly 385
Db      17152 CCAGTTGGTTATGCGGAAAAAGATAATTGGAACTTGTCCGCCAGATGAGTTCAAAGGA 17093
QY      386 ThrPheArgTyrAspGluTyrLeuGluLysGluSerAlaGluThrLeuProLeuAspLeu 405
Db      17092 ACATTCAAGATGAGATGAACTACTTGGAGAAAGAACTCGCAGAAACCTTACCGCTTGA 17033

```

```

QY      406 PheLysProMetProSerGlnGluArgLeuAspLysPheLysVal11eLeu11eSerLys 425
Db      17032 TTCAGCCAAATGCTTCCCAAGAGAGATTAGCAAAATTTAAGTAAATTCGATTTCCAAA 16973
QY      426 Arg----- 426
Db      16972 CGGCTGTTTATATATCGTTTGAGATTGTTTCATTAATAAGTATTCATAATTTGTTCT 16913
QY      427 -----ValGlyLeuArgLeuGluAlaAlaAspMetCysGluAsnGlnPhe11eCys 443
Db      16912 TGTTTAAGTAGAGACTACGCTTGAACTGCACTGACATGCTGGAATAATCAGTTTATTTCT 16853
QY      444 ProAlaThrValLysSerValHisGlyArgLeu11eAsnValAsnPheAspGlyTyrAsp 463
Db      16852 CCAGCTACAGTGAAATTCAGTTTCATGAGAACTGATTAATATGTCATTTGACGCGTGAGAT 16793
QY      464 GluGluPheAspGluLeuTyrAspValAsp----- 473
Db      16792 GAAAGATTTGATGAACGTGTATGATGTGAGAGTCACTTATCATGACCAACGACATTTT 16733
QY      474 -----SerHisAsp11eLeuPro11eGlyTyrCysGluAla 485
Db      16732 CAATGAAATTTCTATCATTTTCAGCTCCCATGATATTTCTACCGATGAGATGTTGTAAGCG 16673
QY      486 HisSerTyrValLeuGlnProProLysLysTyrAsnTyr 498
Db      16672 CACAGTTATGTTCTTCAACCTCCGAAAAAGTAACAATCTAT 16634

```

Search completed: October 28, 2003, 09:42:20  
Job time : 19 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2003, 09:45:58 ; Search time 24 Seconds

(without alignments)  
3.856 Million cell updates/sec

Title: us-09-872-523-6

Perfect score: 1497

Sequence: 1 atgcctgaattcttgaat.....cgaagaagtaactatgta 1497

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 30911 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : 19-apr-1996-271266.\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	1875	85.2	30911	1 CER06C7	ALIGNMENTS ACCESSION:Z71266

RESULT 1  
LOCUS CER06C7/c  
DEFINITION Caenorhabditis elegans cosmid R06C7.  
ACCESSION Z71266  
VERSION GI:1279324  
KEYWORDS adenylsuccinate lyase; myosin heavy chain; Rat TOAD-64 protein  
like; serine/threonine-protein kinase; zinc finger protein.  
SOURCE Caenorhabditis elegans  
ORGANISM Caenorhabditis elegans

Eukaryotes; mitochondrial eukaryotes; Metazoa; Nematoda;  
Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea;  
Rhabditidae; Caenorhabditis.  
1 (bases 1 to 30911)  
Gardner, A.

REFERENCE  
AUTHORS Submitted (19-Apr-1996) Nematode Sequencing Project, Sanger Centre,  
Hinxton, Cambridge CB10 1RO, England and Department of Genetics,  
Washington University, St. Louis, MO 63110, USA. E-mail:  
jess@anger.ac.uk or tw@nemacode.wustl.edu

REFERENCE  
AUTHORS Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berke, M.,  
Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J.,  
Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A.,  
Fulton, L., Gardner, A., Green, P., Hawkes, T., Hillier, L., Jier, M.,  
Johnston, L., Jones, M., Kerhaw, J., Kirsten, J., Laister, N.,  
Latreille, P., Lightning, J., Lloyd, C., McMurray, A., Mortimore, B.,

TITLE  
JOURNAL  
MEDLINE  
COMMENT

O'Callaghan, M., Parsons, J., Percy, C., Rifken, L., Roopra, A.,  
Saunders, D., Showken, R., Smalton, N., Smith, A., Sonhammer, E.,  
Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M.,  
Vaughan, K., Waterston, R., Watson, A., Weinstein, L.,  
Wilkinson-Sproat, J., and Wohldman, P.  
2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
elegans  
Nature 366 (6466), 32-38 (1994)  
94150718  
Current sequence finishing criteria for the C. elegans genome  
sequencing consortium are that all bases are either sequenced  
unambiguously on both strands, or on a single strand with both a  
dye primer and dye terminator reaction, from distinct subclones.  
Exceptions are indicated by an explicit note.  
IMPORTANT: This sequence is NOT necessarily the entire insert of  
clone R06C7. It may be shorter because we only sequence  
overlapping sections once, or longer because we arrange for a small  
overlap between neighbouring subclones.  
The true left end of clone R06C7 is at 1 in this sequence. The true  
left end of clone P21C3 is at 30808 in this sequence. Coding the  
sequences below are predicted from computer analysis, using the  
program GeneFINDER (P. Green, ms in preparation), and other  
available information.  
The end of this sequence (30808..30911) overlaps with the start of  
sequence CEP21C3.

FEATURES  
source

CDS

CDS

```

1..30911
/organism="Caenorhabditis elegans"
/mol_type="unassigned DNA"
/db_xref="taxon:6239"
/chromosome="I"
/clone="R06C7"
complement(join(1315..1480,1528..1715,1747..1783,
1834..2156))
/codon_start=1
/product="R06C7.2"
/protein_id="1279328"
/db_xref="GI:1279328"
/translation="MFSNVPDSDRDKKRYVSTSONYVFLKYSEFNILGKKRIK
NFKLPHAEKRVLPQSVWLKQDTAKTEKRTDITCQSLKLNFEENKRLQDTEREN
KKLNLNLSLEKLKLVSSPENLPLQIFLISENDELREVFHCATFTKLSDFQLF
FPTPMYVANEYAKSGKSNLNEAIIADAIPEVSGPEAARMKOITDSIKXESGFLIDRR
VNASALIIQSNCSIFNSNF"
join(3063..3227,3274..3854,3899..4631,4680..4855,
4899..5074,5117..5463,5510..5749,5790..6097,6143..6254)
/note="protein predicted using GeneFINDER; CDNA EST
yk25b2.5 comes from this gene; CDNA EST yk34a6.5 comes
from this gene; CDNA EST yk31a12.5 comes from this gene;
CDNA EST yk35g4.5 comes from this gene; CDNA EST yk21g1.5
comes from this gene; CDNA EST yk33h1.5 comes from this
gene; CDNA EST yk13a4.5 comes from this gene; CDNA EST
yk13a4.3 comes from this gene; CDNA EST yk21g1.3 comes
from this gene; CDNA EST yk25b2.3 comes from this gene;
CDNA EST yk31a12.3 comes from this gene; CDNA EST yk35g4.3
comes from this gene; CDNA EST yk34a4.3 comes from this
gene"
/codon_start=1
/product="R06C7.1"
/protein_id="1279326"
/db_xref="GI:1279326"
/translation="MSPHPQPHPPMPMPVTPAPGAMTPMPVPVPAQKLOSTGN
DACIKRLQOLNVEDGAKVMKPMPECGMKGRVVDIOTVFGIEVKEKTVHFMHAKA
DUTSTKEVTFPKKREKDPVODRRDKCNIFFLAVEKPEFFKDKGNOIYYDOSTL
YTVNLFSSELDANSTKSRVFOINADTGNDGLKTLPCISLEIVAPRDSITLSENLG
KRTADONIEVNEEYTOFLALNQHCVREINRGCERHGVFLNLTIEGFDORCV
DVGGKOLYPLKKTIOFIEGPGGONNPELVKGGKAAAFHKEQTVIOKLPDITGCP
PSNGLNMTREKKAIVIKGLDVCSTYNNRKHVLIEGIFHSATKTEPLDGTCSI
AEYVADYKKSLSQVPMANLVYCKRGNNNFRPALMYSRQRTIPQOTNSOKKT
KECAVLDVPRKMITTGKNAVNTILELVALDGIKYSSEPLVQAELEDEKELYTOR
SVWSDMKMRAPPEPMFVKPAPVPLMAVAYANGCPSIGDVQVLGMFPGMSTKKIG
MVIKPEPTEGISTEKIMTOLEKVAASKVGLNITDAIVHAKOVKALPQRTMMYV
QDMKISXANAVKDGKRLTENIINKNVKSGGNTYVSDKSKMTBEQLIGVGSVA
PPACTKXKMDKXHLNPQIIGFASNAVNHFEVDVFLAPSGQDTMASIEDVLQNSID

```

**THIS PAGE BLANK (USPTO)**

```
QY 427 GATATGTAATTAATGTAAT----- 447
DB 17871 GATATGTAATTAATGTAATGTAATGTTTTCCTGCAATTAATGTAATATCATCT 17812
QY 448 -----AATGCAATGATGGGAAATCGTCGGCCAAACCTTCGCTGTCTCCAAAT 496
DB 17811 CACAATTCAGAAATTCATTGATGGGAAATCGTCGGCCAAACCTTCGCTGTCTCCAAAT 17752
QY 497 TCGATGAAGGGAAGGCTCTCTTAAGCAACATCGTTTCAAAAGTTGGACAAGCTTTGAAC 556
DB 17751 TCGATGAAGGGAAGGCTCTCTTAAGCAACATCGTTTCAAAAGTTGGACAAGCTTTGAAC 17692
QY 557 TATTAATTAATTCATTAATTCATGAATAGCGGTAGCGGAATTCAGAAATATATGAGC 616
DB 17691 TATTAATTAATTCATTAATTCATGAATAGCGGTAGCGGAATTCAGAAATATATGAGC 17632
QY 617 GACGAATGAATGATCTATCACAAGAAAGACTTTCGCAATCGCTTCAGATGAGATG 676
DB 17631 GACGAATGAATGATCTATCACAAGAAAGACTTTCGCAATCGCTTCAGATGAGATG 17572
QY 677 AGCAGACAAATGCTTTAGCTCTGATCTCAATATTGGATAGACGAGGAAAGCTTCTTCA 736
DB 17571 AGCAGACAAATGCTTTAGCTCTGATCTCAATATTGGATAGACGAGGAAAGCTTCTTCA 17512
QY 737 TATTCCTGTGATTTGAGCAGTCAATGATATCACTAAATGGAAAAAGAAATATA 796
DB 17511 TATTCCTGTGATTTGAGCAGTCAATGATATCACTAAATGGAAAAAGAAATATA 17452
QY 797 TTGACACACAATAATAATTGCTCAAGCAATAAAAAATGAGAAAAATCCAGATATGACT 856
DB 17451 TTGACACACAATAATAATTGCTCAAGCAATAAAAAATGAGAAAAATCCAGATATGACT 17392
QY 857 CAGAGACGTCACATTTGATCAATTAAGCAAAAGATCCAAATTGATCCCATGATTTGAGAA 916
DB 17391 CAGAGACGTCACATTTGATCAATTAAGCAAAAGATCCAAATTGATCCCATGATTTGAGAA 17332
QY 917 AAGTTAAGGTTGACAAAGTTTGAGCTATGACCCCTTGCTCAGCAATTCATAAAC 976
DB 17331 AAGTTAAGGTTGACAAAGTTTGAGCTATGACCCCTTGCTCAGCAATTCATAAAC 17272
QY 977 TCCAGCTGCTCGATTCTCAATTTTGGAAAACTGAAAGATATCTTATTGGGAATGG 1036
DB 17271 TCCAGCTGCTCGATTCTCAATTTTGGAAAACTGAAAGATATCTTATTGGGAATGG 17212
QY 1037 ATGCTCAGATGCACTTGAAACAGTTTCTATTCAATCAATAATACATTTATGTTCC 1096
DB 17211 ATGCTCAGATGCACTTGAAACAGTTTCTATTCAATCAATAATACATTTATGTTCC 17152
QY 1097 CAGTTGTTATGCGGAAAAAGTAAATTTGGAACCTTCTCCGACAGATGCTCAAGAA 1156
DB 17151 CAGTTGTTATGCGGAAAAAGTAAATTTGGAACCTTCTCCGACAGATGCTCAAGAA 17092
QY 1157 CATTGATGAGATGATATCTTGAGAAAGATCTGCAGAAACCTTACCGCTTGACTTGT 1216
DB 17091 CATTGATGAGATGATATCTTGAGAAAGATCTGCAGAAACCTTACCGCTTGACTTGT 17032
QY 1217 TCAAGCCAAATGCTTCCCAAGAGATTAGACAAATTTAAGTAAATCTGATTTCCAAAC 1276
DB 17031 TCAAGCCAAATGCTTCCCAAGAGATTAGACAAATTTAAGTAAATCTGATTTCCAAAC 16972
QY 1277 G----- 1277
DB 16971 GGGTTGTTTAATGCTTTGAGATGTTTCACTAATTAATGTTAATTCATTAATGTTTCTT 16912
QY 1278 -----GGTAGACTACGCTTGAAGCTGTGACATGTTGAAATCAGTTTATTTGTC 1330
DB 16911 GTTTAAGGTAGACTACGCTTGAAGCTGTGACATGTTGAAATCAGTTTATTTGTC 16852
QY 1331 CAGCTACAGTGAATCAGTTTCAATGGAAGCTGATAAATGTCATTTTCAAGGCTGGGATG 1390
DB 16851 CAGCTACAGTGAATCAGTTTCAATGGAAGCTGATAAATGTCATTTTCAAGGCTGGGATG 16792
```

```
QY 1391 AAGATTGATGAACGTGATGATGTA----- 1418
DB 16791 AAGATTGATGAACGTGATGATGATGAGTGAAGTTTATCATGACGAACGACATTTTTC 16732
QY 1419 -----CTCCATGATATTCTTACCATAGATGATGTTGAAAGCC 1456
DB 16731 AATGAAATTTCTATCATTTTCAGCTCCCATGATATTCTTACCATAGATGATGTTGAAAGCC 16672
QY 1457 ACAGTTATGTTCTACCAACCTCGAAAAAGTACAACTATTGA 1497
DB 16671 ACAGTTATGTTCTACCAACCTCGAAAAAGTACAACTATTGA 16631
```

Search completed: October 28, 2003, 09:46:24  
Job time: 26 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 28, 2003, 09:42:00 ; Search time 13 Seconds

(without alignments)  
2.368 Million cell updates/sec

Title: us-09-872-523-5

Perfect score: 2656

Sequence: 1 MSEPFLKIVRANKSDRKLDPK.....PIGWCEAHSYLVQPPKKNY 498

Scoring table:

	BLOSUM62
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Dgapop 6.0 , Dgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1 seqs, 30911 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1 summaries

Command line parameters:

-MODE=frame+pn -DEV=soft -Qus=09-872-523-5 -DB=19-apr-1996-271266  
-SUFFIX=plc -OUT=align5\_1996 -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits  
-START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=1 -DOCALLIGN=200  
-THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=1 -MODE=LOCAL -OUTFMT=plc  
-NORM=ext -HSPALG=500 -MINLEN=0 -MAXLEN=200000000 -NCPU=6 -NO\_XIPXY  
-NEG SCORES=0 -LONGLOG -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : 19-apr-1996-271266:\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	2520.5	94.9	30911	1 CER06C7	ALIGNMENTS ACCESSION:271266

#### ALIGNMENTS

RESULT 1	CER06C7	30911 bp	DNA	linear	INV 19-APR-1996
LOCUS	CER06C7				
DEFINITION	Caenorhabditis elegans cosmid R06C7.				
ACCESSION	271266				
VERSION	271266				
KEYWORDS	G11279324				
SOURCE	adenylosuccinate lyase; myosin heavy chain; Rat TOAD-64 protein like; serine/threonine-protein kinase; zinc finger protein.				
ORGANISM	Caenorhabditis elegans				

Caenorhabditis elegans  
Eukaryotes; mitochondrial eukaryotes; Metazoa; Nematoda;  
Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea;  
Rhabditidae; Caenorhabditis.  
1 (bases 1 to 30911)  
REFERENCE  
AUTHORS  
TITLE  
Direct Submission

#### JOURNAL

Submitted (19-APR-1996) Nematode Sequencing Project, Sanger Centre,  
Hinxton, Cambridge CB10 1RQ, England and Department of Genetics,  
Washington University, St. Louis, MO 63110, USA. E-mail:  
jes@sanger.ac.uk or twanematode.wustl.edu

#### REFERENCE

2 (bases 1 to 30911)  
Wilson, R., Almscough, R., Anderson, K., Baynes, C., Berks, M.,  
Bonfield, J., Burton, M., Dear, S., Du, Z., Duthie, R., Favello, A.,  
Coulson, A., Craxton, M., Connell, M., Cosey, T., Cooper, J.,  
Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M.,  
Johnston, L., Jones, M., Kersey, J., Kirsten, U., Latimer, N.,  
Latreille, P., Lightning, J., Lloyd, C., McMurray, A., Mortimore, B.,  
O'Callaghan, M., Parsons, J., Percy, C., Riken, L., Roopa, A.,  
Saunders, D., Showkeen, R., Smaiden, N., Smith, A., Sonhammer, E.,  
Staden, R., Sultison, J., Thierry-Mieg, J., Thomas, K., Vaudin, M.,  
Vaughan, K., Waterston, R., Watson, A., Weinstock, L.,  
Wilkinson-Sprat, J., and Wohlman, P.

#### TITLE

JOURNAL  
MEDLINE  
COMMENT  
94150718  
Nature 368 (6466), 32-38 (1994)

#### COMMENT

Current sequence finishing criteria for the C. elegans genome  
sequencing consortium are that all bases are either sequenced  
unambiguously on both strands, or on a single strand with both a  
dye primer and dye terminator reaction, from distinct subclones.  
Exceptions are indicated by an explicit note.  
IMPORTANT: This sequence is NOT necessarily the entire insert of  
clone R06C7. It may be shorter because we only arrange for a small  
overlapping sections once, or longer because we arrange for a small  
overlap between neighbouring submissions.  
The true left end of clone R06C7 is at 1 in this sequence. The true  
left end of clone F21C3 is at 30808 in this sequence. Coding  
sequences below are predicted from computer analysis, using the  
program GeneFinder (P. Green, ms in preparation), and other  
available information.  
The end of this sequence (30808..30911) overlaps with the start of  
sequence CEF21C3.

#### FEATURES

source  
1..30911  
/organism="Caenorhabditis elegans"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:6239"  
/chromosome="I"  
/clone="R06C7"  
complement(join(1315..1480,1528..1715,1747..1783,  
1834..2156))  
/codon\_start=1  
/product="R06C7.2"  
/protein\_id="1279328"  
/db\_xref="GI:1279328"

CDs  
join(3063..3227,3274..3854,3899..4631,4680..4855,  
4899..5074,5117..5463,5510..5749,5790..6097,6143..6254)  
/note="Protein predicted using GeneFinder. CDNA EST  
YK252.5 comes from this gene; CDNA EST YK346.5 comes  
from this gene; CDNA EST YK312.5 comes from this gene;  
CDNA EST YK1694.5 comes from this gene; CDNA EST YK2191.5  
comes from this gene; CDNA EST YK93h11.5 comes from this  
gene; CDNA EST YK1344.5 comes from this gene; CDNA EST  
YK134.3 comes from this gene; CDNA EST YK2191.3 comes  
from this gene; CDNA EST YK252.3 comes from this gene;  
CDNA EST YK312.3 comes from this gene; CDNA EST YK1694.3  
comes from this gene; CDNA EST YK1344.3 comes from this  
gene."  
/codon\_start=1  
/product="R06C7.1"  
/protein\_id="1279326"  
/db\_xref="GI:1279326"  
/translation="MSPHPPOPHPPMPMPVTAAGMTMPVPADAKLHSTGN  
DACIRLQQLNVEDGAKMYMKPTBPKMGKRPVDIQTNVGIETVETKETTIVHREMVAKA

DLTSKVEVTFKKGKEDPVQDRBCKNIFFLAVEKNPEFFKMGDGNIVDGGSTL  
YTVVNFSELDNAGTKSVQINGADTGDNDKLTIPCISLETYAPRDNSTLSSNLG  
KRTADQNIENNRNRYTOFLELANOHCAVETNRFGCFEKAQVFLNATEDEQDQCD  
DVGDGKOLYPLGKRTIOFIEGYPGRCONNPSTLVIDMGKAPFKEDQVIOKLDIGOD  
PANGLNMTREKAAAVIKGLDCSYTYNKRRLHIEGI PHEASATKRFPLPGKTSI  
AAYVADKYISLOYPANLVCKDRGNMYPFAELMTSRNORATIPOOTGOSOKT  
KECAVLDPVRORMTIINGKNAVNTLENNLVALGIVYSEPLMVAQRELDGELYOR  
SWMSDMGRAPRPGPFVCPATVPDVAACVNGPGRFSIGVNLQVGFIDSCKKG  
NVIKPECETGLYTEKIMTQLEKVAASKCYLYMTIDAIVLHKQYKALREOTMTIV  
OMKISKANAAYVDGRLTLENINKNTVYKLGNTVYSDAKSMTEBOLIIGVQISA  
PAGTKYMDNKGHLNPOLIGPASNANVHNEFVDPVLAPSGODTMAJEDVLONSID  
LEKNNKALPKRIIVIRSGASGSHSLIAYELPACQALIHGYSKEIKLFIIVKHEH  
SYRFRDQARSGSKATEMNIPRGIVSDANVTPNACKQFNLNHTTIOGAKPLTYTVL  
ADDCAPMDRLEBELFTLCHHOIVSLSTSTPPLVANEVAKRGDLWGEUHTGPIV  
EAKESQGERLKELTEIGYKQDNLQKRVNA"  
complement (join 17895..7963,8094..8340,8392..8457,  
8905..9182,9550..9761,9840..10017,10603..10783,  
10885..10979,11337..11645))  
/note="Similarity to Rat TOAD\_64 protein (PIR Acc. No.  
S49985)"

CDS

/codon\_start=1  
/product="R06C7.3"  
/protein\_id="1279329"  
/db\_xref="GI:1279329"

CDS

CDS

/translation="MWSGHFLKKRDPNVYINLCTLSHAPCTSSPFLTTITIN  
QSPSTQNDFFPFQSLCFVAKIHFNLSNLDSESPVLKNGTVNEDGMPFADV  
LNTVDVCSLLOISCEPNPFLNIWOYTKLRHKEVSPNITLADTEVDATRLVPG  
GIDPHTMOMPMYGEVTKDDPLKGTAAVAGGTMTIDCCDHRNGESLLIAGYRWR  
SWADPKVCCDYGSLVAITWMPETAOMALITSPEGVNSFKFYMAEYETLAVRDEL  
PRGMOECALRALARVHCNGSVIKERIDILAKVTPGEGHQSAPREITEATNRA  
CYLAQANCPYIVHYMTKGASAISSHHAQCSYVGEPIANGALDSGHYNEPMH  
AARVWSPLSPLDPTPELLMKLILAGLELITGTDCYDCKQSLGKNFRTKINGI  
NVEBMSVSWVEKGVSHGIIIDPMRVYSILHSTPAKIFNYPYRKRGLAVSDADIVFN  
PVATRTISKDTHHNDENIFEGINCHGVAEVTISR"  
join(12648..12667,12739..13118,13164..13411,13462..13672)  
/note="protein predicted using GeneFinder: cDNA EST  
yk121911.5 comes from this gene; cDNA EST yk12311.5 comes  
from this gene; cDNA EST yk12449.5 comes from this gene;  
cDNA EST yk59b9.5 comes from this gene; cDNA EST yk12449.3  
comes from this gene; cDNA EST yk59b9.3 comes from this  
gene"

CDS

/codon\_start=1  
/product="R06C7.4"  
/protein\_id="1279327"  
/db\_xref="GI:1279327"  
/translation="MRFPVILIALLLIGASLAHPADPIRAKRDVSAEDEPSSGGEI  
SGSSGSEASGEASGEASGEASGEASGEASGEASGEASGEASGEASGEASGEAS  
VAVDELTLQOLETINETYAQQVAAESQKLIHOANPVYTEMNTLSSANONCILSNIYLA  
NSQMTIDSARLSINETETGTSPATCYSAVCYDSCGSGKCIAGALAGTCNCS  
VFGWPCQEDSACGFGACNSITATDCFAAYTKNLLTAEALTSFCNVETCNGABDN  
VERCHQLPCNYGFCVC"  
complement (join 13796..14018,14067..14234,14299..14462))  
/codon\_start=1  
/product="R06C7.6"  
/protein\_id="1279331"  
/db\_xref="GI:1279331"  
/translation="MISAEALYQSCYCEENYKLEKIDKNGIFAILIISDCKMIP  
LWOKAKSINGVLMWDVHIVIAIEKKNSSKYVLDCTLEWSCDFLEVEKMTNIAEM  
AORDTKFRKFRVIRGENYIISLSDRSHMLSGNGVYLKPPREWLINSHKXSIMEL  
TMSBOIENTYMDPELFOFSK"  
complement (join 114655..114807,14931..15158,15266..16331))  
/note="Similarity to Human adenylsuccinate lyase  
(SW: P088 HUMAN); cDNA EST cm15d4 comes from this gene;  
cDNA EST yk6795.5 comes from this gene; cDNA EST yk6795.3  
comes from this gene"

CDS

/codon\_start=1  
/product="R06C7.5"  
/protein\_id="1279330"  
/db\_xref="GI:1279330"  
/translation="MASEDKFESVLSSTRCKNSPLVLSSETNKATLMROLMIWLAEA  
EKELGKQVTOAIDEMKSNRVDVFWPIRSEERKLRKHVMANHNHAFGLCTPAGIT  
HIGATSCFVQDANADLIAYRDSIDHILKRAIYVIDRAASLCKEYVYVYGRHYOTAS  
LVTVGRGVLMAOELLMAFOSLSEFRDKMRFRGIRKATGTQDSFTLLFAGDSKEVAL

CDS

DELVTKANENRRLITGQRTSRRODSOLVSVSLGLAAKVCYTDIRVLQAFELLE  
PEFKDQIGSSAMPKPKMSEKRCALSRKLIINQOALTLTLAGQERTLDDSGRR  
MLIIPDLVTLAALLTLTQNIIEGSLVQDMADPKKIVBEDIIVGLKEMKMTLEGQVDR  
QOAAHVIRKTLAEKLOLQATQKVIRQVAMPDFDSVDRVAVGLVNNINFTGRCSVO  
TESFLAKLKEPTIDKLYDSAGNVLQDV"  
complement (join 116631..16709,16764..16903,16970..17800,  
17851..17947,18074..18240,18289..18471,18575..18937))  
/note="protein predicted using GeneFinder"

/codon\_start=1  
/product="R06C7.7"  
/protein\_id="1279325"  
/db\_xref="GI:1279325"

/translation="MLKLVILCPALFYNTVSTPFLGCVKQCPDPEFOLTVSHMED  
DGNTPWDEEDITGRMTAFARKTFFYQDGHGFEGLRYGCFPLNCTKNGRREY  
RHGLSSTGSGNGLEYIEYTMSEFLKIVANKKSDRKLDKTYLWESYLHPFKGTSFI  
PVEAFNRLTVNPNECVEGYFETVHVDYDKNDSDJQVRMPARIKVCQGRVLAQFI  
GADRFPMNLISDMPGLANAMSDPMNDKIVVAPLAINEYONDMYVNNCIDGE  
IVGQSLSPKFEDEKALSKRFPVQORLELANSPEIRVARIQETICGRMNVISIT  
KKDPRESIPDADDDROYSSGSOYWDGSEFFIPVGAANGVQANKEVIEHTNK  
IAQAIKNGENRYSDDVTFDQAKDPIIDPMIRKVKVQGFELIDPLAQOPNNIHLVA  
SILFKCTEGYLIVMDGPDALSDSPFIHINTMFPVGAKEYNLDELVPDERKGT  
RWDEYLKESAEPLDLDFKPMPSQERLDKFKVLIISKRVGLRLEADMCNOFICPA  
TVKSYGRLINVFDMDEDEDELVDSDHILPIGWCEASVYLQPKKXNY"  
join(119602..11975,119721..20250,20294..21205,21250..21437,  
21483..22252,22297..22786)  
/note="Similarity to Yeast serine/threonine protein kinase  
BUB1 (SW:BUB1 YEAST); cDNA EST yk6994.5 comes from this  
gene; cDNA EST yk27010.5 comes from this gene; cDNA EST  
yk87910.5 comes from this gene; cDNA EST yk13349.5 comes  
from this gene; cDNA EST yk58d8.5 comes from this gene;  
cDNA EST yk5298.5 comes from this gene; cDNA EST CEBS095R  
comes from this gene; cDNA EST yk27010.3 comes from this  
gene; cDNA EST CEBS095F comes from this gene; cDNA EST  
yk13398.3 comes from this gene; cDNA EST yk58d8.3 comes  
from this gene"

/codon\_start=1  
/product="R06C7.8"  
/protein\_id="1279332"  
/db\_xref="GI:1279332"  
/translation="MSHIRVAFADLPNTSPCTGSETPATQIETLSQMTAEEMKALEY

## Alignment Scores:

Pred. No.:	Length:	30911
Score:	2520.50	Matches: 495
Percent Similarity:	81.08%	Conservative: 2
Best Local Similarity:	80.75%	Mismatches: 1
Query Match:	94.90%	Indels: 116
DB:	1	Gaps: 5

us-09-872-523-5 (1-498) x CER06C7 (1-30911)

Qy	1	MetSerGluPheLeuLeuValAlaArgAlaAsnLysLysSerAspArgLysLeuAspLys	20
Db	18471	ATGTCGAAATTCGAAATTCGAGACTCAACAAAAATCGCAGAAAACTGCATAG	18412
Qy	21	ThrTyrLeuTrpGluSerTyrLeuHisGlnPheGluLysGlyLysThrSerPheIlePro	40
Db	18411	ACCTACTGTCGGAATCTCATTTACATCAGTTCGGAAGAAAACTCTTTCATTC	18352
Qy	41	ValGluAlaPheAsnArgAsnLeuThrValAlaAsnPheAsnGluCysValLysGluLys	60
Db	18351	GTTGAAGATTCATCACTGTAACCTTACATGTAATTTTAAAGAAATGCTAAGAAAGAGTT	18292
Qy	61	Ile-----	PheGluThr 64
Db	18291	ATCGTAGTTCATATTGTCGTAATCGTTTAAATACAATTTTGTAGTTCGAAACA	18232
Qy	65	ValValHisAspTyrAspLysAsnCysAspSerIleGlnValArgTrpPheAlaArgIle	84
Db	18231	GTGGTCATGATTATGACAAAGACCTGCATTCAGATTCAGATGCTTTGCAAGAAAT	18172
Qy	85	GluLysValCysGlyTyrArgValLeuAlaGlnPheIleGlyValAspThrLysPheTrp	104



**THIS PAGE BLANK (USPTO)**





```
CDS
LPEKRAKLPKRIITVRSAGSGSHASILAVIPARAIHHSYKEIKLPIFYVKEH
STRFRDQSRSGKATENMTPPGIVDNAVTPACKQFPLNGHTLLOQAKPLTLVL
ADDCAPMDRELELFTLCHHQIVSLSTSIPTGVANENYAKRRGRDLMGELTTGPI
EAKESQGERLKLKEIGYKQDILNOKRYNA"
complement (join(18955..7963,8094..8340,8392..8457,
8905..9182,9550..9761,9840..10017,10603..10783,
10885..10979,11337..11645))
/notes="Similarity to Rat T0AD_64 protein (PIR Acc. No.
549985)"
/codon_start=1
/product="R06C7.3"
/protein_id="1279329"
/db_xref="GI:1279329"
/translation="WVSSHFLKKDDPNIVNLTGLSSHAPCTSSSFIETITIIIN
OSPSTONODFFEFHQSLCFVFKIHFISSLNDSSESPVINKQTVNEDMDRQAD
LNPVSCSLQICSENPENFLINQWTLRHKVSPNITAPLDEVIDADRLVYIPG
GDPTHMOMVGEVTKDDPLKTEAAVAGGTMIIDCCPDHNGSELILGYNMR
SWADPKVCCDGLSVALITWMPRETBOMAILTSPERGVSFKFMAENTTMVRDEL
PRMGBCATLRLARVHCNSVYKEKEDLAKVGTGEGHTQSRPEIEAATNRA
CVLAQANCPTIVHMTKGAASAISSHQAQSIYFGEPIAGLADGSHYINEMWLH
AARYVMSPLSDPTTPELMLKLAAGELHLTGNTCTYDCQKSLGKGNPKPIINGI
NGVEDRMSVVMKGVSHGIIIDPMRYVSIITSTAAKIFNIYPRKRIIAVQSDADIVFN
PNATRTISKDTHHNLDPNIFEGINCHGVAEVTISR"
join(12648..12687,12739..13118,13164..13411,13462..13672)
/notes="protein predicted using GeneFinder; cDNA EST
yk121911.5 comes from this gene; cDNA EST yk123h1.5 comes
from this gene; cDNA EST yk124a9.5 comes from this gene;
cDNA EST yk59b9.5 comes from this gene; cDNA EST yk124a9.3
comes from this gene; cDNA EST yk59b9.3 comes from this
gene"
/codon_start=1
/product="R06C7.4"
/protein_id="1279327"
/db_xref="GI:1279327"
/translation="MRVFPIALLILITGASLAHPADPIRAKDVASSEDFSGSGEI
SGSSGEASGEASGEASGEASGEASGEASGEASGEASGEASGEASGEASGEAS
VVDDELTLQOELTLNTYAQOVAESKTLHQANPYTTEMTAISANAKNIGISNLYA
NSQWVIDSARLSLNETETETGTSAPTCVSSAVCYGDSGSGKICIGALAGTCNNSC
VFEGPCQEDSACGAGFNACNSITATCDCAAYTKNNLLTAEALTSFCNVTGCAEDN
VEKCHGLPCNYGFCVC"
complement (join(13796..14018,14067..14234,14299..14462))
/codon_start=1
/product="R06C7.6"
/protein_id="1279311"
/db_xref="GI:1279311"
/translation="MISAEALYQSCYCEENVYKLEIKIDKNGIPIAIIISNDCKMIP
LMKOKAKSINGHVLMDHYVIAIEKKNKSKYVDLSTEMSCDFLEVEKTMNLAM
AORDPKFRKFRVIGENYISLSSDRSHMLSKNGVYLKPPBEMULINHKSNNMEL
IMSEIOENTTYMDEBELFOLFSK"
complement (join(14655..14807,14931..15158,15276..16311))
/notes="Similarity to Human adenylsuccinate lyase
(SW: PUR8 HUMAN); cDNA EST cm15d4 comes from this gene;
cDNA EST yk6795.5 comes from this gene; cDNA EST yk6795.3
comes from this gene"
/codon_start=1
/product="R06C7.5"
/protein_id="1279330"
/db_xref="GI:1279330"
/translation="WASEDEKFSFVSTRYCKNSPLVSLISNTKATLTMROLIMILAA
EKEELGKQYTOAIDEMKSNRVPFMRPFRSEBRKIKHVMANHNHAFGLCTAGCIT
HVGATSCFYQDNADILAYRDSIDHILKRAIVYIDRAASLKNKEVYVGRTHYOTAS
LVTYGRGVLAQOELMAFQSLSEFIDKRAIVYIDRAASLKNKEVYVGRTHYOTAS
DELVTKKANSFRLITQOTYGRQOQDSQVLSLGAARQVCDIDRLQAQFGEILR
PEKQDIGSSAMPYKKNPKMSRCCSKRLINAPBEALITLADQCLERTLDSARL
MLIPDLITAEALITLIONIPGLISVQTONVKKIVYDELAIFGLGAEAMNLEEGDR
QQAHAVIRKTALEAKQLOATQKYDIDQTMADPFPDSVRRVGVVANNPINFGRQCYSO
TSEFTAKELKPTIDKYLDSAGNVQLDV"
complement (join(16631..16709,16764..16903,16970..17800,
17851..17947,18074..18240,18289..18471,18575..18937))
/notes="protein predicted using GeneFinder"
/codon_start=1
/product="R06C7.7"
/protein_id="1279325"

CDS
```

```
/db_xref="GI:1279325"
/translation="MLKLVITICRLFYNTVTSSTRFLPCVRYKCPDVEYFQLTVSHMED
DGNFPMWRDEITGRMTMFAKKIFFYQDGHGGEFGLEFYGVFLHCKNKNFREY
RGHSJSTSGSDGSEIETRYMSEFLAKIIRANKSKSPDKTYLMSYLYHQEKRYTSFI
PVEAFENMLTYNVEVCEVGEIETVYVNDKQDSIQVRFARIEKQGRVLAQFI
GADRFKPMNLISDMFGPAAAMADPMNDKIVAPPLAINEEYONDMVYNNVNCIDGE
IVGQTSLSPRKDEKALLSKRFRVQGRLELNTNSNRIIVARIIOELICGRMVSIT
KKQDPRESIPDADROQVSSQYVWIDBSFFITPVGFAANGVQANAKKXIVHTNK
IAQAIKGENENRVDSDVTFDQAKDPIDEPHTRKRVQGFELIDPAQOPNULHVA
SLKFRCTEGYLIVMDGPDLAEDSPFHLINNTMFWPVAEKYNLELVPDERKGTFR
RMDEYLEKESAEETLPDLFKPMPSQERLDKRVKILSRVGLRLEALDMCENQFCIPA
TVKSHVGRLLINVFNDGMBDEDELVDYDSHLLIGICEALSYVLOPKKXNY"
join(19602..19675,19721..20250,20294..21205,21250..21437,
21483..22252,22297..22786)
/notes="Similarity to Yeast serine/threonine protein kinase
BUB1 (SW: BUB1 YEAST); cDNA EST yk6944.5 comes from this
gene; cDNA EST yk27c10.5 comes from this gene; cDNA EST
yk87g10.5 comes from this gene; cDNA EST yk133g8.5 comes
from this gene; cDNA EST yk58d8.5 comes from this gene;
cDNA EST yk5298.5 comes from this gene; cDNA EST CEB5095R
comes from this gene; cDNA EST yk27c10.3 comes from this
gene; cDNA EST CEB5095F comes from this gene; cDNA EST
yk133g8.3 comes from this gene; cDNA EST yk58d8.3 comes
from this gene"
/codon_start=1
/product="R06C7.8"
/protein_id="1279332"
/db_xref="GI:1279332"
/translation="MSHIRVAFAPLDTPNPTSCGLETFATQIETLSQMTABEMKAEYL

Query Match      85.2%; Score 1275; DB 1; Length 30911;
Best Local Similarity 81.3%; Pred. 0;
Matches 1497; Conservative 0; Mismatches 0; Indels 344; Gaps 5;

QY      1  ATGCTGAATTCGAAATATGTCAGAGCTAACAAAAATCGACAGAAATCTGCATAG 60
DB      18471  ATGCTGAATTCGAAATATGTCAGAGCTAACAAAAATCGACAGAAATCTGCATAG 18412

QY      61  AACTACTGTGGGAATCCTATTATACAGTTCGAGAAAGAAAACTTCTTTCATTC 120
DB      18411  AACTACTGTGGGAATCCTATTATACAGTTCGAGAAAGAAAACTTCTTTCATTC 18352

QY      121  GTTGAAGATTCATCGTAACCTTAAGTAAATTTTAAACGATGCGTAAGAGAGATT 180
DB      18351  GTTGAAGATTCATCGTAACCTTAAGTAAATTTTAAACGATGCGTAAGAGAGATT 18292

QY      181  ATC-----TTGAAACA 192
DB      18291  ATC-----TTGAAACA 18232

QY      193  GTGTCATGATTATGACAGAACTGCGATTTCAGTTCAGTTCAGTTCAGTTCAGT 252
DB      18231  GTGTCATGATTATGACAGAACTGCGATTTCAGTTCAGTTCAGTTCAGTTCAGT 18172

QY      253  GAAAAAGTTGGGATACAGAGTCTGCGTCAAGTTATCGAGAGCGAACAAATTTGG 312
DB      18171  GAAAAAGTTGGGATACAGAGTCTGCGTCAAGTTATCGAGAGCGAACAAATTTGG 18112

QY      313  CTCATATTTTATCGAGAGATATGTTTGGTTGGCAAA----- 350
DB      18111  CTCATATTTTATCGAGAGATATGTTTGGTTGGCAAA----- 18052

QY      351  ----- 350
DB      18051  TCTACTATTCGAATAAATAATGTTCTGTTACATAAATTTCTAGAGAACATCGTATTA 17992

QY      351  -----CGCGGCAATGCGTAT 366
DB      17991  AAACTTCGAACATTTGTATATAGTAAATTTGAAACATTTGAGCGCGCAATGCGTAT 17932

QY      367  CCCAATATGATTAATAATTTGATATGCTCGCGCGCTTGCAATCAAGAGAAATACCA 426
DB      17931  CCCAATATGATTAATAATTTGATATGCTCGCGCGCTTGCAATCAAGAGAAATACCA 17872
```



**THIS PAGE BLANK (USPTO)**